AD-A051 142   NAVAL OCEAN SYSTEMS CENTER SAN DIEGO CA   F/G 9/2

MODELS OF OPTIMAL FILE ALLOCATION IN A DISTRIBUTED DATA BASE: A--ETC(U)

JAN 78 T H CROCKER, D M KLAMER

UNCLASSIFIED   NOSC/TR-195   NL

1 OF 1
ADA
051142

END
DATE
FILMED
4 -78
DDC

(12)

NOSC

NOSC TR 195

NOSC TR 195

Technical Report 195

MODELS OF OPTIMAL FILE ALLOCATION IN A DISTRIBUTED DATA BASE: A SURVEY

Classifies distributed file models and describes the parameters needed in modeling

TH Crocker
DM Klamer

15 January 1978

Prepared For
Naval Electronic Systems Command
(ELEX 330)

Approved for public release; distribution is unlimited

NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CALIFORNIA 92152

NAVAL OCEAN SYSTEMS CENTER, SAN DIEGO, CA 92152

AN ACTIVITY OF THE NAVAL MATERIAL COMMAND

RR GAVAZZI, CAPT, USN
Commander

HL BLOOD
Technical Director

ADMINISTRATIVE INFORMATION

This report was performed under NOSC problem CC2582, and was sponsored by Naval Electronic Systems Command, ELEX 330. The work described herein was completed by the Tactical Command Control Division, Code 8242, Naval Ocean Systems Center.

Released by
RC Kolb, Head
Tactical Command Control
Division

Under authority of
JH Maynard, Head
Command Control – Electronic
Warfare Systems and Technology
Department

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE (14)		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NOSC Technical Report 195 (TR 195)	2. GOVT ACCESSION NO. NOSC/TR-195	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) MODELS OF OPTIMAL FILE ALLOCATION IN A DISTRIBUTED DATA BASE: A SURVEY. Classifies distributed file allocation models and describes the parameters needed in modeling.		5. TYPE OF REPORT & PERIOD COVERED (9) Technical rept. 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) TH Crocker [redacted] DM Klamer		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Ocean Systems Center San Diego, California 92152		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NOSC CC2582
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Electronic Systems Command (ELEX 330)	(11)	12. REPORT DATE 15 Jan[redacted]78 (12) 13. NUMBER OF PAGES ~~28~~ 30p.
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

ARPANET
Data bases
Models

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Examines the existing distributed data base file allocation models and gives a breakdown of the models by type (deterministic one-phase, deterministic multi-phase, stochastic discrete, stochastic continuous). The relationships and identities used to describe the models are divided into four categories: file information and parameters, transmission characteristics, computer characteristics, and costs. In the investigations which led to this paper it was seen that the models defined were initially very general. The models included relationships which were very detailed in their description of the file allocation problem. In previous analyses using these models, simplifications were often made for computational tractability. Many of the assumptions and models ended up so restricted in scope or detail as to be unrealistic. There is a great need for more work in this area.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102 LF 014 6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

393159

OBJECTIVE

Investigate the problem of minimal cost allocation of files in a computer network with respect to the design of efficient computer networks; establish the relationships which need to be considered to accurately model the file allocation problem in a computer network.

RESULTS

1. Existing distributed data base file allocation models are broken down by type (deterministic one-phase, deterministic multi-phase, stochastic discrete, stochastic continuous) and reviewed. They are described in terms of file information and parameters, transmission characteristics, computer characteristics, and costs.

2. The models defined are found to be initially very general, but simplification for computational tractability results in restricting them so severely in scope or detail that they end as unrealistic.

RECOMMENDATION

1. Develop and analyze realistic adaptive stochastic models.

CONTENTS

I. INTRODUCTION

The use of distributed data bases became a practical reality with the implementation of the ARPANET. Before that time the trend in computing had been toward centralized computing resources where the power of a large machine was deemed economical. Concomitants of decentralization are decreased communication costs, increased efficiency due to computer specialization, and increased system reliability due to system redundancy. With a centralized computer system all remote users must communicate with the computer for all computer interactions, while with a computer network, users may do most of their communications with their local computer, only occasionally communicating with remote computers for data or programs. Though computers can do many tasks well, some are more efficient at particular tasks than others. Some computers may perform scientific calculations very efficiently while others may perform input/output operations very efficiently. Currently, no computer performs all operations optimally. Thus a computer network can give a user access to a machine which can handle his particular problem most efficiently, while in a centralized computer system the mainframe may be required to perform many tasks for which it is not well suited. System reliability is enhanced with a computer network because the system does not depend on just the operation of one computer complex. If one computer system should go down most users still have access to the other computing resources of the network, while in a centralized computing system if the computer goes down the whole system is unavailable to all the users. For these reasons, as well as others, computer networks will be used in more and more applications, in particular, military applications.

Military systems seem particularly suitable for implementation in a distributed computer network. In military systems it is very important to communicate information with superior, collateral and subordinate commands. It is also important that each command have computing power available to it. Computing power local to the command to control equipment, such as radars, guns and missiles, is needed at each command. Much of the information collected by commands is shared with the other commands. Currently, a great portion of this communication is done by teletype messages, a slow medium. The information contained in the messages is not necessarily tailored to the needs of the unit receiving the messages as it might be if it were the response to a query. Computer to computer communication would allow less duplication of information, transmittal of more relevant information and quicker transmission and usage of the information. For these reasons the study and eventual efficient implementation of military distributed data bases are necessary.

The objective of our work is the design of efficient computer networks. To pursue this objective we are currently investigating the problem of minimal cost allocation of files in a computer network. In this report we review much of the previous work in this area to establish the relationships which need to be considered to accurately model the file allocation problem in a computer network. Most of the previous work has assumed complete knowledge of the data base parameters. Solutions to the minimal cost allocation problem were fixed in time, either in a one-time period or multi-time period problem. In our research, of which this paper is the first step, we want to consider the dynamic adaptive allocation of files.

The remainder of this report is divided into five sections which lead to the cataloguing of relationships which will be used in modeling adaptive distributed data bases. Section 2 classifies the distributed data base allocation work into four types of models: deterministic one-phase, deterministic multi-phase, stochastic discrete, and stochastic continuous. In Section 3, examples of the models of the above types are given, and the relationships and assumptions used to define them are detailed. In Section 4, the work done by previous authors is critiqued. In Section 5, a list of relationships is compiled which will be used to model command control distributed data base systems. A summary of this work is given in Section 6.

II. DISTRIBUTED DATA BASE MODEL CLASSIFICATION

Models which have been used for describing distributed data bases can be classified into two primary groups. The two primary groups, deterministic and stochastic, can be further divided into two subgroups each. The four classification groups for distributed data base models are:

(i) Deterministic – one-phase

(ii) Deterministic – multi-phase

(iii) Stochastic – discrete time

(iv) Stochastic – continuous time

The classification is based upon model assumptions and does not necessarily reflect the true environment of the distributed data base which is being analyzed.

The distinction between deterministic and stochastic models is as follows. In a deterministic model all of the relevant information is assumed to be known ahead of time. Such quantities as (1) the probability that a user will request a specific file, (2) the rate at which a file is used, (3) the change of file usage patterns, etc, are assumed to be known prior to system design. In addition, the model parameters do not change unless the new value of the parameter and the time when it changes are known. On the other hand, stochastic models allow for unknown parameters. These parameters are usually estimated and then the distributed data base dynamically reconfigures to optimize system performance.

Deterministic systems are categorized into one-phase deterministic systems and multi-phase periodic deterministic systems. In a one-phase deterministic model all parameters are assumed to be known *and constant*. System performance is optimized for the fixed parameters and no changes are made to the data base after this initial design. A multi-phase periodic deterministic system allows changes to occur, however, these variations must be known exactly. The data base system may also be assumed periodic. For example, a day may be broken into an 8-hour day shift and a night shift. File usage rates would change when the shifts change and the times of the shift change would have to be known in addition to knowing the rate changes. As can be seen from the general description of these deterministic models, the models are not flexible and, in general, are not realistic models to describe a distributed data base in the real world.

Stochastic models allow for some parameters to vary or to be unknown. In a discrete time stochastic model the system is monitored at a sequence of times, and is then reorganized based upon estimates and other available information to optimize the system performance. An example of this situation is when the distributed data base under consideration is linked together by asynchronous communication lines. In a continuous time stochastic model, events are allowed to occur at any time. For example, a file usage rate may be unknown and continuously varying with time. Note that stochastic models allow for a changing environment which is not known *a priori*. Thus, it seems that stochastic models provide a natural environment for describing distributed data bases.

III. DISTRIBUTED DATA BASE MODELS – EXAMPLES

In this section a brief summary of the models used to describe the file allocation problem in distributed data bases is given. The models will be described chronologically within the breakdown given in the previous section.

Deterministic – One-phase

The original work in file allocation was by Chu.[1] He considered the following zero-one programming model. Consider a network with $i = 1, \ldots, n$ computers and $j = 1, \ldots, m$ files. Let X_{ij} indicate the j^{th} file is stored on the i^{th} computer,

$$X_{ij} = \begin{cases} 1 & j^{th} \text{ file stored in } i^{th} \text{ computer} \\ 0 & \text{otherwise} \end{cases}$$

Let r_j be the number of redundant copies of file j. The following constraint is required

$$\sum_{i=1}^{n} X_{ij} = r_j$$

Let L_j be the length of the j^{th} file and b_i be the available memory size of the i^{th} computer. Then the memory constraint implies

$$\sum_{j=1}^{n} X_{ij} L_j \leqslant b_i$$

Let a_{ijk} denote the expected time for the i^{th} computer to receive the j^{th} file from the k^{th} computer. Let T_{ij} be the maximum expected allowable retrieval time of the j^{th} file to the i^{th} computer. Then a_{ijk} must be less than T_{ij}.

$$(1 - X_{ij}) X_{kj} \, a_{ijk} \leqslant T_{ij}$$

Note that if $r_j = 1$ then $X_{ij} X_{kj} = 0$ for $i \neq k$ and the above equation reduces to

$$X_{kj} \, a_{ijk} \leqslant T_{ij}$$

In many of the models that follow it is assumed that a_{ijk} is a constant, but in this case a_{ijk} is estimated as a function of the file access rate and line transmission speed. The structure of a_{ijk} assumes that

$$a_{ijk} = w_{ik}^{(1)} + t_{kj} + w_{ki}^{(2)}$$

where $w_{ik}^{(1)}$ is the expected queuing delay at the i^{th} computer for the channel to the k^{th} computer, $w_{ki}^{(2)}$ is the expected queuing delay at the k^{th} computer for this channel to the

i^{th} computer, and t_{kj} is the expected computer access time to the j^{th} file. The superscripts indicate priority classes of messages. To simplify the analysis it is assumed that t_{kj} is small compared to the queuing delays, and, further, that the delay of the short high priority request messages is much shorter than the longer low priority reply messages. Thus a_{ijk} is approximated by $w_{ki}^{(2)}$ using a queuing model with Poisson arrivals of requests at a rate λ_{ik} between the i^{th} and k^{th} computer. To complete this queuing model the following parameters are needed: l_j the length of the file portion sent in response to the query. Note $l_j \leq L_j$. The average time to transmit the reply from the k^{th} to the i^{th} computer is $1/\mu_{ik}$. The request rate for the j^{th} file at the i^{th} computer is u_{ij}.
Then

$$\lambda_{ik} = \sum_{j=1}^{m} u_{ij} (1-X_{ij}) X_{kj}$$

Let

$$1/\mu_j = l_j/R$$

where R is the transmission rate from the k^{th} to the i^{th} computer. Then $1/\mu_j$ is the transmission time. The average time required to transmit a reply is then

$$1/\mu_{ik} = (1/\lambda_{ik}) \sum_{j=1}^{m} u_{ij} (1-X_{ij}) X_{kj}/\mu_j$$

The traffic intensity is defined to be

$$\rho_{ik} = \lambda_{ik}/\mu_{ik} = \sum_{j=1}^{m} u_{ij}(1-X_{ij}) X_{kj}/\mu_j$$

Then queuing theory for a Poisson arrival, constant service time model implies the waiting time

$$w_{ki}^{(2)} = \rho_{ik}/(2\mu_{ik}(1-\rho_{ik})) \quad \text{for } i \neq k$$

Taking the formulas above, combining them with the restriction $w_{ki}^{(2)} \leq T_{ij}$ implies

$$(1-X_{ij}) X_{kj} \lambda_{ik} - 2\mu_{ik} (\mu_{ik} - \lambda_{ik}) T_{ik} \leq 0$$

The above define all the constraints in Chu's model. All that needs to be defined now is the objective function. The cost of the model to be minimized involves storage and transmission costs,

$$C = C_{storage} + C_{transmission}$$

where

$$C_{storage} = \sum_{i,j} C_{ij} L_j X_{ij}$$

and

$$C_{transmission} = \sum_{i,j,k} C'_{ik} \ell_j u_{ij} X_{kj} (1-X_{ij})/r_j + \sum_{i,j,k} C'_{ik} \ell_j u_{ij} X_{kj} P_{ij} .$$

The first term in the transmission cost represents costs due to file request and the second term represents costs due to file update where P_{ij} is the frequency of modification of file j on computer i. For multiple copies of files this model is not linear, but since the programming problem is a zero-one problem there are standard methods for linearizing the model which requires additional constraints. The model is not very efficient. For a simplified three-computer, five file, one-copy network the time to find the optimum solution was 25 seconds on an IBM 360/65.

Whitney[2] considers the following problem which is of the one-phase deterministic variety. Let G(T,L) be the graph of the network on which the distributed data base is located, where T is the set of nodes which represents the locations in the network where the terminal users are located and L represents the edges of the graph which are communication lines between user nodes. Communication costs are given as weights $\{g_m | m \epsilon L\}$ on the edges of the graph. Let S_{ij} represent the weight of shortest path from node i to node j, ie S_{ij} represents the cost of the least expensive communication route from computer site i to site j. Given the graph G(T,L), then an efficient algorithm developed by Hu determines the path which yields the minimum communication cost between the two nodes of interest. This minimum communication cost, S_{ij}, is given as

$$S_{ij} = \min_{P_{ij}} \sum g_m, \text{ where } P_{ij} \text{ is a path from node i to node j.}$$

The next quantity considered in this model is the message traffic from user terminal j to file k, which is represented as $\gamma_j(k)$ and given by

$$\gamma_j(k) = P_j(k)\, R(k)$$

where $P_j(k)$ is the probability that file k is requested from site j and R(k) is the rate at which some record of file k is requested. All records of a file are homogeneous, and, therefore, R(k) represents the rate of request for each record of file k. The quantity $\gamma_j(k)$ corresponds to the rate of message traffic for records of file k which is generated by user terminal i. Now the cost to minimize for the assignment of file k is

$$\min_{i\epsilon T} \sum_{j\epsilon T} \gamma_j(k)\, S_{ji}$$

The cost of assigning file k to site j is given by

$$t(k,j) = \sum_{i\epsilon T} \gamma_i(k)\, S_{ij}$$

and is a minimum when

$$t(k,j) \leqslant t(k,i)$$

for every $i\epsilon T$. Although not explicitly mentioned by Whitney, it appears that each of the costs $t(k,j)$, $j\epsilon T$, must be computed and then a comparison of the costs must be done in order that the minimum cost be found. The model is also generalized slightly by allowing multiple copies of a file to exist in the data base.

The following paper by Casey[3] presents a simpler model than the above models in most ways except that it treats the number of copies of each file as a variable. Casey assumes the files are independent of each other and thus may optimize one file at a time. Three costs are considered: storage, query, and update. The query and update costs consist of communication costs. Thus if I is an index set representing the computers on which a given file resides the cost C(I) is given by

$$C(I) = \sum_{j=1}^{n} \left(\sum_{k\epsilon I} \psi_j\, d'_{jk} + \lambda_j \min_{k\epsilon I} d_{jk} \right) + \sum_{k\epsilon I} \sigma_k$$

where

σ_k is the storage cost at the k^{th} computer

d_{jk} is the cost of communication from node j to node k for a query

d'_{jk} is the cost of communication from node j to node k for an update

λ_j is the volume of query traffic emanating from node j

ψ_j is the volume of update traffic emanating from node j

Note that $C(\phi)=\infty$ so that the problem is nontrivial. Using this model Casey proves some theorems about the optimal number of copies of files in the network and properties of the optimal file allocation.

A later paper by Morgan and Levin[4] considers a model more general than Casey's and different from Chu's. Their model considers both programs and data files. They differentiate between programs and files because they assume that in a heterogeneous network

files would be transferable among all the computers but programs would be transferable only among homogeneous subsets of computers. In their model, described below, the indices i,j,k refer to computers, f to files and p to programs. The network is assumed to have N computers, F files and P programs. Queries and updates are assumed to be processed through programs. Let

λ_{ipf} be query traffic from node i to file f via program p

λ'_{ipf} be update traffic from node i to file f via program p

C_{ij} be communication cost per query unit from i to j

C'_{ij} be communication cost per update unit from i to j

σ_{jf} be storage cost of file f at j

σ'_{jp} be storage cost of program p at j

α be expansion factor for query message

β be expansion factor for update message

J_p be set of nodes where program p may be stored

The expansion factors are the ratios of the length of query (update) from a program to the length of query (update) from the originating node. To define the model the following control variables are required:

$$y_{kf} = \begin{cases} 1 & \text{if copy of file f is stored at node k} \\ 0 & \text{otherwise} \end{cases}$$

$$y'_{jp} = \begin{cases} 1 & \text{if copy of program p is stored at node j} \\ 0 & \text{otherwise} \end{cases}$$

$$x_{jkf} = \begin{cases} 1 & \text{if transactions from node j to file f are routed to node k} \\ 0 & \text{otherwise} \end{cases}$$

$$x^{f}_{ikp} = \begin{cases} 1 & \text{if transactions from node i to file f are routed to node j via program p} \\ 0 & \text{otherwise} \end{cases}$$

The following two parameters define traffic flow:

$$\rho_{jf} = \sum_{i,p} \lambda_{ipf}\, x^{f}_{ijp} \quad \text{is query traffic to file f processed at node j}$$

$$\psi_{jf} = \sum_{i,p} \lambda'_{ipf} \quad \text{is update traffic to file f processed at node j}$$

The model may now be described. The objective is to minimize

$$C = \sum_{f,j,i,p} \lambda_{ipf} * C_{ij} * x^{f}_{ijp}$$ = Communication cost of queries from initiating nodes to the programs

$$+ \sum_{f,j,i,p} \lambda'_{ipf} * C'_{ij} * x_{ijp}^{f}$$ = Communication cost of updates from initiating nodes to the programs

$$+ \sum_{f,j,k} \rho_{if} * \alpha C_{jk} * x_{jkf}$$ = Communication cost of queries from programs to files

$$+ \sum_{f,j,k} \psi_{if} * \beta * C'_{jk} * y_{kf}$$ = Communication cost of updates from programs to files

$$+ \sum_{f,k} \sigma_{kf} * y_{kf}$$ = Storage cost of files

$$+ \sum_{j,p} \sigma'_{jp} * y'_{jp}$$ = Storage cost of programs

Subject to the following constraints:

- To assure the attainment of a feasible solution there must be at least one copy of each file and each program,

$$\sum_{j} y'_{jp} \geq 1 \qquad p=1, \ldots, P$$

$$\sum_{k} y_{kf} \geq 1 \qquad f=1, \ldots, F$$

- To assure that every transaction to every file, via every program and from every node, will have a defined route:

$$\sum_{j} x_{ijp}^{f} \geq 1 \qquad i=1, \ldots, N; p=1, \ldots, P; f=1, \ldots, F$$

$$\sum_{k} x_{jkf} \geq 1 \qquad j=1, \ldots, N; f=1, \ldots, F$$

- To assure residency of the appropriate files and programs in accordance with the defined routes

$$\sum_i x_{ijp}^{f} \leq N\, y'_{jp} \qquad j=1, \ldots, N;\ p=1, \ldots, P;\ f=1, \ldots, F$$

$$\sum_j x_{jkf} \leq N\, y_{kf} \qquad k=1, \ldots, N;\ f=1, \ldots, F.$$

- To assure that program p will reside only in a node at which it can be processed

$$y'_{jp} = 0 \qquad j \notin J_p, \quad p = 1 \ldots P$$

And y'_{jp}, y_{kf}, x_{ijp}^{f}, x_{jkf} are binary variables.

Deterministic – Multi-phase

Levin and Morgan[5] also consider deterministic multi-phase models. Their model is a generalization of their one-phase model. Using the decomposition result mentioned above they define a model to minimize the cost of allocating one file at a time. The model above is used with the following changes:

$$y_{kt} = \begin{cases} 1 & \text{if file copy is assigned to node k at period t} \\ 0 & \text{otherwise} \end{cases}$$

$$K_t = \left\{k \mid y_{kt} = 1\right\} \text{ is an arbitrary assignment of a file at period t}$$

$$\underline{K}_T = \left\{K_1, \ldots, K_T\right\} \text{is an arbitrary arrangement of file assignments at periods 1 to T}$$

The cost consists of two parts. First is the sum of the operating costs $C(K_t)$ over all time periods given by the model in the previous section. The second is the transition costs to change the file allocation from one period to the succeeding period. To determine this cost a little additional notation is necessary. Let L_s denote the length of the file in storage units. Let γ be the transformation factor from storage units to message units. Then the number of message units, L_m, required to transmit the file is given by

$$L_m = L_s \gamma$$

Then the transition cost from period t-1 to period t is given by

$$T_t(K_{t-1}, K_t) = \sum_{t \epsilon K_t} L_m \min_{\ell \epsilon K_{t-1}} C_{\ell k}$$

This cost is determined by sending the file over the most economical path. Combining the above equations the total cost over T time periods is given by

$$G(\underline{K}_T) = \sum_{t=1}^{T} C(K_t) + T_t\,(K_{t-1}, K_t)$$

The optimal solution $\underline{K}_T^*$ is given by

$$G(K_T^*) = \min_{\underline{K}_T} G(\underline{K}_T)$$

The final multi-phase deterministic model to be discussed is by Segall.[6] Segall extends the scope of this model from deterministic to adaptive models to be described in a later section. To obtain the adaptive results, a model which is much more restrictive than the model of Levin and Morgan is considered. The following general assumptions are made in this model; there is only one copy of each file in the network, files are short so that there are no storage or communication limitations, the files are requested according to a mutually independent process, and files may only move from computer to computer in response to a request. These assumptions imply the files may be considered independently of each other. With these assumptions the model may now be defined for a network of two computers. Let

$$Y_i(t) = \begin{cases} 1 & \text{if file stored at computer i at time t} \\ 0 & \text{otherwise} \end{cases}$$

where $i = 1,2$ and $t = 1,2,3, \ldots$ and

$$\{n_i(t) \;:\; t = 1,2,\ldots\}$$

be two independent binary sequences describing the file requests at each computer, where $n_i(t) = 1$ if there is a request for the file from computer i at time t and $n_i(t) = 0$ otherwise. Let $a_i(t)$ be the random rate of $n_i(t)$. Let $B(t)$ contain all past information relevant to the evolution of $n(\cdot)$.

Then

$$P\{n_i(t) = 1 \;:\; B(t-1)\} = a_i(t)$$

Let C_{ij} denote the cost of transmitting the file from computer i to computer j. Let C_i denote the cost of storing the file at computer i. Then the expected cost over all time periods is

$$C = E\sum_{t=1}^{T}\left\{\sum_{i=1}^{2}\left(C_i Y_i(t) + \sum_{j=1}^{2} C_{ij}\, Y_i(t)\, n_j(t)\right)\right\}.$$

The control variables of the process are

$$u_i(t) = \begin{cases} 1 & \text{if decision is to transfer file to memory } i \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

The evolution of the process is defined by

$$Y_1(t+1) = Y_1(t)\,(1-u_2(t)) + Y_2(t)\,u_1(t)$$

$$Y_2(t+1) = Y_1(t)\,u_2(t) + Y_2(t)\,(1-u_2(t)).$$

Since the files may only move in response to a request the controls are functions of the $Y_i(\cdot)$. Thus

$$u_1(t) = \alpha_{21}(t)\,Y_2(t)\,n_1(t)$$

$$u_2(t) = \alpha_{12}(t)\,Y_1(t)\,n_2(t).$$

The optimization problem is then to find the control laws $\alpha_{ij}(\cdot)$ and initial locations to minimize the expected cost of the network operation. For deterministic models Segall also considers the following deterministic continuous time model, because in a network with more than two computers the analysis is easier. Let

$$\{N_i(t) \;:\; t > 0\} \qquad i = 1, 2, \ldots, M$$

be M independent counting processes representing the file requests at the individual computers. Let $\lambda_i(t)$ be the random rate of the requests. The costs C_i and C_{ij} are the same as in the previous paragraph. Thus over a period of time T the total expected costs would be

$$C = E\left\{\int_0^T \left[\sum_{i=1}^{M} C_i Y_i(t)\,dt + \sum_{i=1}^{M} \sum_{j\neq i} C_{ij}\,Y_i(t-)\,dN_j(t)\right]\right\}.$$

The controls of the process are defined to be

$$\alpha_{ij}(t) = \begin{cases} 1 & \text{file in computer } i \text{ at time } t- \text{ and requested by computer } j\neq 0 \text{ at time } t \\ & \text{and decision is to erase at } i \text{ and store at } j \\ 0 & \text{otherwise.} \end{cases}$$

The dynamics of the files are given by

$$dY_i(t) = -Y_i(t-)\sum_{j\neq i}\alpha_{ij}(t)\,dN_j(t) + \sum_{j\neq i}\alpha_{ji}(t)\,Y_j(t-)\,dN_i(t).$$

The problem is then to find the optimal controls $\alpha_{ij}(\cdot)$ which minimize the operation cost of the model.

Stochastic – Discrete Time

There are two types of stochastic discrete models. The first is by Levin and Morgan[5] which is an extension of their previously described deterministic models. The second is by Segall[6] which, though it has more restrictive assumptions, is broader in scope. This will be clarified below.

In the model of Levin and Morgan they assume the request rate λ_{ipt} and update rate λ'_{ipt} are random variables. Rather than optimizing $G(\underline{K}_T^*)$ Levin and Morgan optimize

$$EG(\underline{K}_T^*) = \min_{\underline{K}_T} EG(\underline{K}_T)$$

This is equivalent to optimizing the original model with the request and update replaced by the expected values of the request and update rate. To obtain the optimal allocation, the request and update rates are estimated statistically and substituted for the expected values of the request and update rates. The optimal allocation is then determined as in the deterministic cases.

The paper by Segall gives an adaptive file allocation algorithm as opposed to the static allocation algorithm of Levin and Morgan. This is an extension of his two-computer deterministic models given above. The notation used here will be that given above. It is assumed that only one of the time varying rates is random. The system dynamics are

$$Y_1(t+1) = Y_1(t)\,[1-u_2(t)] + Y_2(t)\,u_1(t)$$

$$Y_2(t+1) = Y_1(t)\,u_2(t) + Y_2(t)\,[1-u_1(t)]$$

with the controls

$$u_1(t) = \alpha_{21}(t)\,Y_2(t)\,n_1(t)$$

$$u_2(t) = \alpha_{12}(t)\,Y_1(t)\,n_2(t)$$

where, here it is assumed, $n_2(t)$ is Bernoulli with known rate $a_2(t)$ and $n_1(t)$ has a random rate $a_1(t)$. The random rate $a_1(t)$ will be modeled as a finite-state Markov process with states

$$\rho^{(1)} < \rho^{(2)} < \cdots < \rho^{(m)},$$

transition probabilities

$$\Pr\{a_1(t+1) = \rho^{(j)} \mid a_1(t) = \rho^{(i)}\} = q_{ij}(t),$$

and initial distribution

$$\Pr\{a_1(1) = \rho^{(k)}\} = \pi_k, \quad k = 1,2, \ldots, m.$$

Define the variable

$$x_k(t) = \begin{cases} 1 & \text{if } a_1(t) = \rho^{(k)} \\ 0 & \text{otherwise} \end{cases}$$

Let $\underline{x}(t)$ be the vector with the components x_i, $i=1, \ldots, m$. Let $\hat{\underline{x}}(t/t-1)$ be the least squares estimate of $\underline{x}(t)$ given $\{n_1(1), \ldots, n_1(t-1)\}$. Then the problem is to find the optimal controls to minimize the cost function

$$C = E\sum_{t=1}^{T}\{[C_1 + C_{12}\, a_2(t)]\; Y_1(t) + [C_2 + C_{21}\, \hat{a}_1(t/t-1)]\; (1-Y_1(t))\}$$

where $\hat{a}_1(t/t-1)$ is the estimate of $a_1(t)$ given information up to time t.

Stochastic – Continuous Time

As in the discrete case Segall[6] models the continuous time request rates as a finite-state Markov process. The notation will be the same as the above notation for the continuous time deterministic multi-phase model. The dynamics of the file are again given by

$$dY_i(t) = -Y_i(t-)\sum_{j\neq i}\alpha_{ij}(t)\; dN_j(t) + \sum_{j\neq i}\alpha_{ji}(t)\; Y_j(t-)\; dN_i(t).$$

The problem is to determine the dynamic controls

$$\alpha_{ij}(t) = \alpha_{ij}^{*}\,(t, N_i(s), Y_i(s), s < t, i = 1, \ldots, M)$$

that minimize the cost

$$C = E\left\{\int_0^T\left[\sum_{i=1}^{M} C_i Y_i(t)\; dt + \sum_{i=1}^{M}\sum_{j\neq i} C_{ij}\; Y_i(t-)\; dN_j(t)\right]\right\}.$$

Defining

$$J(t) = i \text{ if } Y_i(t) = 1,$$

there is a one to one correspondence between $J(\cdot)$ and $Y(\cdot)$, thus the cost may be written more compactly as

$$C = E\int_0^T L'\,(t, J(t), \hat{\underline{\lambda}}_1(t))\; dt.$$

To avoid notational difficulties Segall further assumes that only $\lambda_i(t)$ is random and that $\lambda_i(t)$, $i=2, \ldots, M$ are deterministic. The random rate $\lambda_1(\cdot)$ is modeled as a finite-state Markov process with states

$$\rho^{(1)} < \rho^{(2)} < \ldots < \rho^{(m)}$$

and transition probabilities

$$\Pr\{\lambda_1(t+dt) = \rho^{(j)} \mid \lambda_1(t) = \rho^{(k)}\} = q_{kj}(t)\,dt + 0\,(dt)\quad k \neq j$$

$$\Pr\{\lambda_1(t+dt) = \rho^{(k)} \mid \lambda_1(t) = \rho^{(k)}\} = 1 + q_{kk}(t)\,dt + 0\,(dt)$$

where

$$q_{kk}(t) = -\sum_{j \neq k} q_{kj}(t)$$

and initial distribution

$$\Pr\{\lambda_1(0) = \rho^{(k)}\} = \pi_k.$$

Define

$$x_k(t) = \begin{cases} 1 & \text{if } \lambda_1(t) = \rho^{(k)} \\ 0 & \text{otherwise.} \end{cases}$$

Then if $\hat{\underline{x}}(t)$ is the least squares estimate of

$$\underline{x}_k(t) = [x_1(t), \ldots, x_m(t)]^T$$

and

$$\underline{\rho} = [\rho^{(1)}, \ldots, \rho^{(m)}]^T$$

then the best estimate $\hat{\lambda}_1(t)$ of $\lambda_1(t)$ given $\{N_1(s), s \leqslant t\}$

is

$$\hat{\lambda}_1(t) = \underline{\rho}^T \hat{\underline{x}}(t).$$

Then the cost may be written

$$C = E \int_0^T L(t, J(t), \hat{x}(t))\,dt$$

where $\hat{\underline{x}}(t)$ satisfies the m-dimensional recursive equation

$$d\hat{\underline{x}}(t) = [Q^T(t)\hat{\underline{x}}(t) - P(\hat{\underline{x}}(t))\underline{\rho}]\,dt + \frac{P(\hat{\underline{x}}(t-))\underline{\rho}}{\underline{\rho}^T \hat{x}(t-)}\,dN_1(t)$$

where

$$Q(t) = [q_{kj}(t)] \quad k, j = 1, 2, \ldots, m$$

and

$$P(\hat{\underline{x}}(t)) = E^{\mathcal{F}t}\,[\hat{\underline{x}}(t)^T\,\hat{\underline{x}}(t)].$$

To solve this model results from the theory of stochastic processes are required.

IV. CRITIQUE OF MODELS

The first research in the allocation of files in a distributed data base is due to Chu.[1] Chu treats the deterministic one-phase model. His work has the most general assumptions of all the papers considered herein. As the papers became more current the models became less general. For example, Chu is the only author to consider a finite memory limitation. However, the later papers consider more general types of problems, for example stochastic or multi-phase models. In this section a review of each of the papers by category will be given.

Deterministic One-phase

Most of the authors consider models of this type. As mentioned above Chu's model is the most general. He considers limited memory, file length, communication delay as a function of the file request rates, update rates, storage costs, communication costs, and multiple copies. These factors are used to build a zero-one programming problem which can be transferred, using standard techniques, into a linear zero-one programming problem. This model laid the framework for the succeeding models. The principal drawback of Chu's model is that it requires that all the parameters, eg number of copies of files, request rates, etc, be known and constant throughout the use of the system. Another drawback is that the algorithm for optimizing the performance of this model is computationally very slow.

Whitney's model of file allocation is only part of an overall system design model for a network. The solution technique used to allocate the files is not very elegant; the cost of assigning a file to each site must be computed and then the site associated with the minimum cost is allocated the file. In order to compute the cost of assigning a file to a particular site, all of the routes between the proposed site and sites which request the file must be enumerated. In any system with a large number of nodes the allocation of just a single file would be a formidable task to perform because of the problem of enumerating all the possible routes between two nodes. In addition to the enumeration problem all parameters associated with a file (request rates from each user, length, etc) must be known prior to system design.

Casey, noting that Chu assumes the number of copies of files to be known, attacks the problem of determining the optimal number of copies of files in a distributed data base. Though his model is not as general as Chu's, it considers the costs of locating files at nodes, communication costs of queries and updates and volume of communication and update traffic. This model allows each file to be treated independently. Using these assumptions Casey is able to prove theorems which determine the optimal number of copies of files

and procedures for efficiently determining the locations of the files. The principal drawbacks in this work are the limited assumptions, independence of files and unlimited memory, which are too simplified to be realistic.

The major contribution of Levin and Morgan is in the extension of the deterministic one-phase model into multi-phase and stochastic models, to be discussed below. In the one-phase deterministic model they differentiate between programs and data files. This is done for two reasons. First, programs are not necessarily compatible with all computers in a network, while data files can be made compatible. Second, programs can initiate requests for other files while the reverse does not occur. This model also has no memory limitations. The model is more general than Casey's but not more general than Chu's. The communication costs are given and are not a function of the message traffic. These assumptions allowed Levin to prove that the file allocation problem can be decomposed to individual file minimization problems. The optimal file allocation can be obtained by optimizing the location of one file at a time. Such results are interesting and can extend the intuition with regard to file allocation. But assumptions which neglect the interrelationships of files are lacking in realism.

The models discussed above are all deterministic one-phase models. This type of model assumes that all the characteristics of the network are known at design time and that they remain the same thereafter. This assumption is a severe restriction. The model described in the next section is one generalization.

Deterministic Multi-phase

Levin and Morgan generalize their one-phase model into a multi-phase model. The model is deterministic in that the assumption is made that for each time period all the characteristics of the model are known. The same basic model is considered as in the one-phase model with the addition of a transition cost between time periods. This takes care of the cost of transferring files from their allocations at one time period to the next. The authors do not give an algorithm for solving this problem. They do refer to another paper in which a dynamic programming solution is discussed. The major drawback in this model is that the system designer must know the file usage in advance, as in the one-phase model. Realistic systems of the future will necessarily have to be adaptive and deterministic models will only be useful for either preliminary analyses or finding approximate solutions.

Stochastic Discrete

The stochastic model of Levin and Morgan is based on their multi-phase deterministic model. The equations are the same. The only difference is that they assume the request

and update rates are random variables rather than known quantities. Thus in order to optimize the objective function they must consider the expected value of the deterministic objective function. Levin[7] demonstrates that the optimal solution to the file assignment reduces to that of estimating the first moment of the access rates distributions. The drawback to this work is that the optimal solution is not adaptive but requires estimation of the parameters at system design time.

The paper by Segall, though it has much more restrictive assumptions (considers one file at a time, fixed communication costs and only one copy of file in the network), gives a model whose solution to the file assignment problem is adaptive. Segall develops a finite-state Markov process and uses dynamic programming to obtain the optimal solution. This paper is the only paper to date which obtains an adaptive solution to the file assignment problem.

Stochastic Continuous

In the discrete case Segall solved the problem for a two-computer network. For a computer network with more than two computers there is a finite probability that more than one computer will request the file at a given time. In a continuous model the probability of that event is zero, which simplifies the analysis. Segall again for the stochastic continuous model derives an adaptive solution. The principal drawback to this work is that the models are too simple to be of practical use. But the work is important because it gives an adaptive solution to the file optimization problem in a distributed data base.

V. RELATIONSHIPS USED IN MODELING DISTRIBUTED DATA BASES

This section contains lists of assumptions, parameters, costs and relationships that different authors have considered when describing models of distributed data bases and computer networks. These lists provide a summary of pertinent ideas when developing a model which describes a distributed data base in a computer network. The four primary categories considered when modeling a distributed data base are: (1) file information and parameters, (2) transmission characteristics, (3) computer characteristics, and (4) costs. Some overlap exists between the lists.

File Information and Parameters

The salient features which describe files for the purpose of setting up a model of a distributed data base are the number of copies of each file, the length of each file, and the rate or frequency at which each of the files is accessed. These features are listed in Table I.

The number of copies of individual files is usually assumed to be known at the time of system design. In addition, most models assume that only one copy of a particular file exists in the distributed data base since the analysis and modeling are simpler than in the case of multiple copies. Another option to choose from when setting up the model is to let the number of copies of an individual file be a variable used in the optimization procedure.

For the length of the file one of two choices is usually made. The length of a file is assumed to be known or the length of the files assumed to be short. When the length of a file is known, memory restrictions are placed on each of the nodes in the distributed data base and the available memory at each node is a restriction placed on the model for optimization. The use of short files arises from several assumptions. One can argue that the cost of memory is inexpensive and, therefore, the amount of storage needed at each node is an irrelevant factor to consider when optimizing the system performance. The fact that storage capacity at computer sites may be fully utilized places this argument upon untenable grounds. Another argument is more direct and to the point; the files are assumed to be short because this implies that the files are independent of each other and, hence, no interaction between the files takes place—not to mention the fact that the analysis of the optimization problem is easier to accomplish.

Query and update rates are other file parameters which are considered when modeling a distributed data base. For some models query (request for information only) rates and update (change the contents of the file only) rates are combined into a single request rate. Most models assume that the rates are known prior to system design. When the rates are not known, then an estimate must be formed in real time and the system, possibly, has to reconfigure itself based upon the estimated information.

Transmission Characteristics

The objective of considering transmission characteristics is to determine the cost of transferring information between nodes of the distributed data base and ensure a rapid response to queries and updates. The transmission characteristics are summarized in Table II. One of the concerns with transmission characteristics between nodes is the time to retrieve and transfer a file from one node of a network to another node. In addition, constraints or priorities may be placed upon the transmission channel, or the messages (file queries and updates) and file transfers may be modeled as a queuing system.

The model structure for determining the cost of transferring information between nodes can be simple or elaborate. The simplest model for transmission cost is to lump all the cost into one quantity which represents the cost to transmit a particular file from one node to another node. Constraints, such as maximum retrieval and transfer time and transmission channel capacity, are placed on some of the models. At the extreme of the message transfer model are elaborate message queuing structures which take into account the average query and update rates of a particular file, random lengths of messages, priorities on different types of messages and information transfers, and average message traffic between nodes.

Computer Characteristics

Computer characteristics, listed in Table III, which are relevant to modeling distributed data bases for the purpose of optimal file allocation are the amount of memory, file access and retrieval time, and the compatibility of programs and files on different machines. The primary computer characteristic considered is the available memory at a site for accepting the transfer of a file. As mentioned above, this option of limited memory may, or may not, be invoked.

Another computer characteristic which is considered in some models is the access time required to retrieve a file from a particular storage medium, such as disk access time as opposed to tape drive access time. However, after breaking the cost (or time) of obtaining access to file into fine detail, eg query time plus disk access time plus reply time, an assumption is generally made to simplify the analysis. An example of a simplifying assumption which one model uses is that the machine access time is much shorter than either the query or response message time and, therefore, the disk access time can be ignored altogether.

Another alternative is to allow data to be transferred to any node of the network while programs can be transferred only to a limited subset of the network nodes. Two points should be noted here. First, the term file can be used in a generalized sense to include a

program, as well as data files. Hence, all of the models considered could be used for both program and file allocation in a network. Second, restricting the nodes of a network to which a file (program or data file) can be transferred is usually a simple constraint which could be placed on most of the models without increasing the complexity of analysis. In models which assume that the files are independent, restricting the allowable nodes to which a file can be transferred imposes no additional constraints, however, for other models this restriction may be nontrivial.

Costs

The final category of relationships to be considered when developing a model for file allocation in a distributed data base is a collection of miscellaneous costs (Table IV). These costs include storage costs, query and update costs, reconfiguration costs, and communication costs.

The only costs which all models consider are the cost of storing a file at a particular node location and the transmission cost of sending the information of one file from one node to another node. In the case of storage costs of a file at a particular site, a cost is charged for the storage of a file regardless of the memory restriction. In particular, if the model assumes no restriction on memory, a cost is given to the amount of storage that a file requires. As with message transmittal the transmission costs may be simple or complex and, in fact, the costs for transmission are based directly upon the modeling of the message transmittals. The communications cost can be divided into query cost and update cost. The query cost can be further divided into a query (interrogation) cost and response cost, along the same lines as mentioned above.

The cost of reconfiguration or transition when a file is moved to a new site is necessary for the deterministic multi-phase and stochastic models. This transition cost includes only the cost of sending the file from one node of the data base to another node. No overhead cost for reconfiguration has been provided in the transition cost.

The relationships discussed above describe the factors which have been considered in modeling distributed data bases. Optimal allocation of files in a network is a fairly new field of research and the relationships listed should not be considered to exhaust the factors which might be needed to model realistic distributed data bases.

TABLE I. FILE INFORMATION AND PARAMETERS

1. Number of copies of a particular file
   a. Given at design time
   b. Variable – one of the parameters to be used in the optimization procedure
2. Length of file
   a. Short – no interaction between files
   b. Known length
3. Request rates for information contained within the file
   a. Rate at which a particular program requests a file
   b. Rate at which a node in the network requests a file
4. Update rates for modifying a file
5. Query rates for obtaining information
6. File dependence
   a. Independent of each other – no interaction between the files
   b. Dependent upon each other

TABLE II. TRANSMISSION CHARACTERISTICS

1. Time to retrieve file from one node of the network to another node
2. Maximum retrieval time
3. Transmission channel capacity
4. Message queuing
   a. Average delay in sending request or query
   b. Average delay in receiving a reply from a query after it has been sent
   c. Poisson arrivals of messages – request and reply
5. Random lengths of messages
6. Priorities
   a. Short requests – high priority
   b. Long replies – low priority
7. Rate of message traffic from file to user

TABLE III. COMPUTER CHARACTERISTICS

1. Memory
   a. Finite amount of memory
   b. No restriction on memory
2. File update and retrieval time
3. Programs only run on specific machines

TABLE IV. COSTS

1. Storage cost
   a. File storage cost
   b. Program storage cost
2. Communications cost
3. Query cost
4. Update cost
5. Reconfiguration or transition cost when a file moves to a new site
6. Communication cost of queries
   a. Program (user) to file
   b. File to user communication costs
7. Communication cost of updates (same as 6a and b)

VI. SUMMARY

In this paper we have examined the existing distributed data base file allocation models. A breakdown of the models by type (deterministic one-phase, deterministic multiphase, stochastic discrete, stochastic continuous) was given. The relationships and identities used to describe the models were divided into four categories: file information and parameters, transmission characteristics, computer characteristics, and costs. In the investigations which led to this paper it was seen that the models defined were initially very general. The models included relationships which were very detailed in their description of the file allocation problem. In previous analyses using these models, simplifications were often made for computational tractability. Many of the assumptions and models ended up so restricted in scope or detail as to be unrealistic. There is a great need for more work in this area.

REFERENCES

1. WW Chu, "Optimal File Allocation in a Computer Network," Computer Comm Networks, Prentice Hall 1973, p 82-94

2. V Whitney, A Study of Optimal File Site Assignment ..., PhD Thesis, Univ of Mich (1970)

3. RG Casey, "Allocation of Copies of a File in an Information Network," Spring Joint Computer Conference 1972, p 617-625

4. HL Morgan and KD Levin, "Optimal Program and Data Locations in Computer Networks," CACM Vol 20 (1977), p 315-21

5. KD Levin and HL Morgan, "Optimizing Distributed Data Bases – a Framework for Research," National Computer Conference 1975, p 473-478

6. A Segall, "Dynamic File Assignment in a Computer Network," IEEE TAC, Vol 21 (1976), p 161-173

7. KD Levin, Organizing Distributed Data Bases in Computer Networks, Penn Univ Decision Sci Dept, Rept 74-09-01, 1974